

Canonical Workflow for Experimental Research

Dirk Betz^{1,2†}, Claudia Biniossek^{1,2}, Christophe Blanchi³, Felix Henninger⁴, Thomas Lauer², Philipp Wieder⁵, Peter Wittenburg^{3,6} & Martin Zünkeler⁷

¹TIB – Leibniz Information Centre for Science and Technology, Welfengarten 1 B, 30167 Hannover, Germany

²Center for Empirical Research in Economics and Behavioral Sciences (CEREB), University of Erfurt, Nordhäuser Str. 63, 99089 Erfurt, Germany

³DONA Foundation, C/O Université de Genève, Rue du Général-Dufour 24, 1204 Genève, Switzerland

⁴Statistics and Data Science in Social Sciences and the Humanities, Department of Statistics, Ludwig-Maximilians-University München, Ludwigstr. 33, 80539 München, Germany

⁵GWDG, Burckhardtweg 4, 37077 Göttingen, Germany

⁶FDO Forum, Gemeindeweg 55, 47533 Kleve, Germany

⁷KAIROS Bochum, Universitätsstraße 136, 44799 Bochum, Germany

chinaXiv:202211.00448v1

Keywords: Experiments; Experimental economics; FAIR digital objects; Information science; FAIR research output

Citation: Betz, D., et al.: Canonical workflow for experimental research. *Data Intelligence* 4(2), 155-172 (2022). doi: 10.1162/dint_a_00123

Received: September 18, 2021; Revised: December 17, 2021; Accepted: March 1, 2022

ABSTRACT

The overall expectation of introducing Canonical Workflow for Experimental Research and FAIR digital objects (FDOs) can be summarised as reducing the gap between workflow technology and research practices to make experimental work more efficient and improve FAIRness without adding administrative load on the researchers. In this document, we will describe, with the help of an example, how CWFR could work in detail and improve research procedures. We have chosen the example of “experiments with human subjects” which stretches from planning an experiment to storing the collected data in a repository. While we focus on experiments with human subjects, we are convinced that CWFR can be applied to many other data generation processes based on experiments. The main challenge is to identify repeating patterns in existing research practices that can be abstracted to create CWFR. In this document, we will include detailed examples from different disciplines to demonstrate that CWFR can be implemented without violating specific disciplinary or methodological requirements. We do not claim to be comprehensive in all aspects, since these examples are meant to prove the concept of CWFR.

[†] Corresponding author: Dirk Betz (Email: Dirk.Betz@tib.eu; ORCID: 0000-0002-6411-4758).

1. INTRODUCTION

Research based on experiments in controlled environments is one of the fundamental scientific paradigms to gain new insights. In this paper, we will focus on experimental research with human subjects in different areas such as psycholinguistics, psychology, economics, the social sciences, medicine, and related research areas. In the realm of a large German funding program^① experts from these areas came together to discuss the characteristics of their experimental processes in detail. The goal was to identify commonalities and create a shared perspective on improving FAIRness [1]. The key observations from this discussion confirmed the findings from a comprehensive analysis of many research infrastructure projects [2]:

- There are only marginal differences when a specific experimental paradigm (e.g., randomized controlled trials, game theory experiments measuring behaviour, factorial surveys measuring attitudes) is used across disciplines, but substantial differences (e.g., in ontologies, metadata, and processing) between different methodological procedures. It is important to note that it is not the domains and disciplines (e.g., Sociology) but the methodological paradigms (survey experiments, psychological experiments in the laboratory or field, behavioral game theory experiments in the laboratory, field, or online) that yield the specifics, such as finding appropriate ontologies and metadata to describe and document the experiment. For example, all behavioral game theory experiments can be described within the same metadata schema, regardless of whether they were conducted in Sociology, Political Science, or Economics.
- The FAIR principles were known to all experts; however, actual research practices hardly changed to implement these principles to the majority of studies with often relatively small (monetary) resources and/or communities with a lack of adequate research data infrastructures and best practices. One reason for this procedural inertia might come from the fact that extensive adaptations of various tools would be necessary to realize a general implementation of the FAIR principles into the existing research workflows. At the same time limited developer capacities and the constant need for prioritized methodological features, restrict the resources for these adaptations.
- There are many repetitive steps in the daily work starting with the preparation of an experiment to analysing collected data. Yet no workflow mechanisms are in place to cover these repetitions and allow researchers to benefit from more efficient processes.

The group of experts agreed that (1) major steps and investments would be necessary to change the situation towards more efficiency and a higher degree of FAIRness, (2) a joint approach across disciplines and across methodological paradigms would be reasonable given the similarity of workflows, and (3) practices are required that take the load off from the researchers to create the necessary motivation to comply with these workflows. This agreement was made in full consciousness of the fact that the adaptations of existing solutions may be expensive and, in some cases, even impossible. Since in most experimental labs^② specific sequences of actions need to be taken repeatedly, the implementation of canonical workflows

^① <https://www.nfdi.de/>

^② We use the term “lab” in a broad meaning for all kinds of data-generating, managing, and processing places including field and online-experiments.

consisting of harmonised components without adding load on the researchers seems a promising approach. A widely accepted solution for low-effort documentation is to immediately create Fair Digital Objects^③ (FDOs)^④ [3, 4] at each workflow step. FDOs “bind all critical information about an entity in one place and create a new kind of actionable, meaningful and technology independent object.” (Source: <https://fairdo.org/>). This approach would also allow embedding existing tools. In the first instance, wrappers could be used to integrate these tools where possible until the tools would have been changed to support FDOs.

The overall expectation of introducing CWFR and FDOs can be summarised as:

- reducing the gap between workflow technology and research practices to make research processes based on experiments more efficient,
- improving FAIRness without adding administrative load on the researchers by, e.g., automatically generated metadata and archiving of the digital object.

In the following section, we will present nine distinct and atomic steps that can be identified as sufficiently similar in experiments with human subjects across disciplines. Some of these steps can be skipped when implementing the workflow if they are not required by institutional or legal regulations.

The analysis of data collected from an experiment will not be considered in this paper as an integral part of the basic experiment CWFR for two reasons. First, the process of analysing data is considerably more complex than all other steps within the workflow and would create a bottleneck when it comes to identifying similarities. Second, data analysis is typically much more differentiated across disciplines and woven much deeper into well-established practices. Including this step would likely hamper any harmonisation attempt. We are, however, convinced that the data analysis with a different sequence of steps requiring different strategies and different components could be turned into a separate canonical workflow. It should also be noted that despite all harmonisations in many experimental paradigms that suggest using a workflow framework there will always be studies requiring specific developments and a particular set of actions.

2. A COMMON EXPERIMENTAL WORKFLOW PATTERN

In Figure 1 we present the canonical experiment workflow consisting of nine atomic steps that cover the typical process of a research project that is based on a controlled experiment. The preparation phase includes five steps and starts with discussing the intentions and expectations. This needs to result in a hypothesis which needs to be formulated in some form dependent on the lab environment. In well-organised labs, this will be described in the form of a short document. Following this step, a suitable experimental design will be chosen which is largely a prose text which may include references to so-called experimental

^③ <https://fairdo.org/>

^④ It should be noted that this paper is not meant to discuss architectural issues. Here we refer to the documents about the Fair Digital Objects.

paradigms often used for classes of experiments in the different disciplines and sometimes references to experiment execution software which is suitable to carry out this work. After iterations, this step will result in a short document.



Figure 1. An abstract representation of experimental workflows is indicated with a set of recurring canonical actions.

The next step includes the specification of detailed experimental parameters such as type and set of stimuli to be presented, concrete actions to be taken during the experiment execution, timing of the actions, types, and number of subjects to be included in the experiment to come to reliable results. To select the stimuli and the subjects, this step often includes the use of specific databases, software, and tools that contain pre-programmed stimuli or the participant pool. Each lab has its own way of organising and structuring this kind of data and thus is using specific databases and tools, and access to the subject database might be restricted due to protecting personal data. Also, the list of parameters to be specified depends on the experimental software being used which suggests using formal key-value pairs in the specification document so that later transformers can be easily built. To select stimuli and subjects it may be necessary to split this step up in a short sequence of steps each associated with a specific software component.

In many research fields and institutions, it is required to ask for an ethics review by an institutional review board^⑤. While the formats for the different institutions might be different, the set of information to be provided can mostly be generated from the already created descriptions, i.e., if the information from the first step is structured it is possible to create the ethical requests with the help of simple transformers. In some domains, simplified ethics reviews have been realized^⑥ and could lead to an actionable ethics review request management tool. In some cases, researchers do preregistration of their experiments which also can be generated based on the information that has already been entered.

^⑤ It should be noted that in the case of an ethics review in the medical field, automatization is hardly possible due to legal regulations, but the “Gesellschaft für experimentelle Wirtschaftsforschung (GfEW)” has already set up a simplified ethics review process that appears to be automatable. The Erfurt Lab is also currently pushing ahead with the development of a (partially) automated ethics review. The goal is now to (a) automate each workflow step as much as possible and (b) create a FAIR Digital Object (FDO) at each step, so that, unlike in the past, there is no need to laboriously document, curate, and archive the data after the manuscript has been submitted to a publisher.

^⑥ see <https://gfew.de/en-ethik>

The ethical review can raise some concerns that may lead to a revision of the experimental setup or some components therein. In other cases, the results of a pre-test might create the necessity to re-adjust some of the experimental parameters. When the ethical request was positively reviewed, and the tests were positive the actual experiment can begin. It needs to be mentioned here that in general special software is being used to support tests and experiment execution. The rationale behind this is that these experiment programs are highly special since specific hardware needs to be interfaced, special tight timing constraints need to be respected, and a particular sequence of actions has to already be implemented, etc. This implies that a wrapper needs to be developed that extracts all parameters needed from the existing data structures, be transformed into the required structures, and then submitted as attributes in the call to start the experiment software. Often stimuli are presented as lists in a separate file, i.e., one of the attributes will be a reference to a list.

chinaXiv:202211.00448v1

In general, experimental actions are repeated over M items and N subjects, i.e., a micro-workflow embedded in the experiment software is repeated for $N \times M$ times which will include some measurements. Therefore, the result file in general includes $N \times M$ vectors with all relevant information necessary for the analysis. Experiments with many subjects, however, are not being executed in one run. Often these experiments are stretched over long periods and, in some cases, the same subjects are being tested again after some time, e.g., for longitudinal experimental studies. In such situations, the data that is collected needs to be integrated into one data set ready to be analysed. Such data sets are then being stored and registered. To prevent data loss or the accidental exclusion of sessions, different measures need to be taken which could also help in preventing "fraud". Backup copies of each session will be necessary but will not be sufficient as long as hashing of the content is not incorporated in some way as it is suggested by using FDOs.

As shown in Figure 1 the canonical workflow framework will help the researcher from the beginning to create proper documentation which will then be extended at each step so that at the end a comprehensive metadata description will be ready without the need to retype information. This functionality is often part of virtual research environments (VRE) designed and configured to facilitate typical tasks related to project administration and research data management.

In these workflows, some specificities need to be taken care of.

Revision

In some steps, revisions within the action may occur. If they are not implemented by "micro-workflows" within a specific software, the canonical workflow needs to have facilities to handle these revisions. As an example, we could refer to tests that might have to be done several times with some adjustments of parameters, which in general require human interaction.

Non-Linearity, Splitting, and Merging

At each step the researcher must be able to go back to a previous step and start with the descriptions made beforehand by adapting the existing descriptions, i.e., no retyping is required. This implies that at any step a Digital Object must be created that contains all information and points to other information, e.g., referring to prior FDOs containing former versions.

Researchers might also want to start some steps in parallel, for example, after the experimental design they immediately want to start the ethical request procedure. Parallel actions will be started and finished asynchronously. In addition, a timer may be started to remind the researcher to check for the state of the ethical review process. Facilities need to be provided to merge these different paths again.

Researchers often conduct more than one experiment at a time to test various experimental settings, i.e., parallelism needs to be taken care of at the project level in addition to parallelism within the workflow. Every project runs its own instance of the framework.

Interoperability

CWFR is an excellent vehicle to improve FAIRness and thus interoperability of the data produced substantially^②. Currently, most data is being exchanged without any further metadata associated with it. This results in the well-known 80% loss in efficiency since the data receivers need to find out what the data is about, how it can be interpreted, etc. CWFR has the chance to make metadata creation for the researchers as easy as possible since it will guide them from the beginning of the experiments and allow them to stepwise add information by automatic measures such as extracting header information from recordings. In addition, by introducing FDOs as underlying mechanisms the strong relation between PID, data, and metadata (of different sorts) will not be lost over time. It is the strength of FDOs to keep the binding between all this relevant information as long as needed. CWFR based on FDO technology is the way to introduce FAIR compliance and increase interoperability without adding additional load on the researchers.

3. FINAL CANONICAL WORKFLOW AND VIRTUAL RESEARCH ENVIRONMENTS

Concerning workflows as sketched above two major phases can be distinguished:

- Phase 1 is characterised by specifying the sequence of actions to be carried out and the type of data/metadata needed.
- Phase 2 is characterised by executing such a specified sequence of actions.

As already indicated, in this paper we will not discuss the analysis phase since different characteristics and types of processes are involved. We will also not discuss interactive workflow frameworks separately, since the specification and the execution steps are combined in one framework associated with asynchronous processes.

3.1 Preparation Phase

The experimental project begins with creating a first empty metadata FDO called Exp_MD_FDO. A simple editor can begin to add the description of the intentions, hypothesis, and the usual entries such as researcher name, experiment name, date, etc. to this FDO. For later extraction purposes it would be helpful

^② It should be noted that having FAIR data is not sufficient for interoperability, but a necessary step.

to structure the input by using agreed keywords. The next step using an editor is to describe the experiment design which is a more prose-like description which can be used, for example, for the ethical review, in publications, etc. Already at an early-stage researchers ask for ethical permission to carry out the intended experiment. If necessary, an ethical request is packaged from the already specified information in Exp_MD_FDO and sent to the corresponding board by email. In the case of structured information in Exp_MD_FDO, this process of mapping information to given templates can be automated.

In parallel, researchers start defining the parameters of their experiments which is highly dependent on the experimental paradigm and the chosen software. To optimally support the user, it would be useful to start an editor and to invoke a template that is associated with the software and the paradigm being used. As indicated above, this step may involve some actions that are widely taken care of by the experimenter. The set of micro-actions[®] in the experiment needs to be defined and their timing, the set of stimuli per micro-action needs to be selected and probably ordered, the set of subjects that will participate in this experiment will be selected which mostly involves email interaction and many adjustments.

Mostly special software is being used to access the pool of stimuli (acoustic, visual, etc.) and to make the selection as efficient as possible. The design informs how many such stimuli targets are required and which characteristics they need to have. This selection will result in an ordered list of values and/or paths to files which can be used by the experiment execution software.

Also, in this case mostly special software is being used to access the subject database and to select the set of appropriate subjects needed for a specific experiment. A list of possible candidates is being prepared, much interaction needs to be done to check availability and suitable dates and finally, subjects are requested to participate in the experiments at certain dates & times. Often subjects are identified by specific codes only known to a responsible subject pool manager and these codes are then transferred to the execution software to include them in the result files.

In this paper, we assume that both lists, stimuli, and subjects^{®,} are being stored as DOs for documentation purposes, and that the Exp_MD_FDO will include the PIDs of these two FDOs which we will call Exp_Stim_FDO and Exp_Subj_DO. There is no doubt that suitable tools need to be available to make these transitions as simple as possible for the researcher. Storing them as FDOs and DOs will also allow experimenters to reuse them for other purposes in time with small changes, i.e., the time-consuming selection process can be reduced. As medicine is one of the most regulated areas, it might serve as a stress test at this workflow step for Exp_Subj_DOs, adhering to the following applicable quality guidelines/standards: GCP—Good Clinical Practice Guide Requirements for electronic trial data handling systems; GAMP 5 Guide: Compliant GxP Computerized Systems Computerized system validation of GxP systems; ISO 13485: 2016—Medical devices—Requirements for regulatory purposes.

[®] In this paper, we use the term micro-action to refer to all detailed steps that need to be carried out to run an experiment which is in general covered by specialised software.

[®] It should be noted that mechanisms need to be available to anonymize subject names before continuing usage. Most experimental software packages already collect data in an anonymized way.

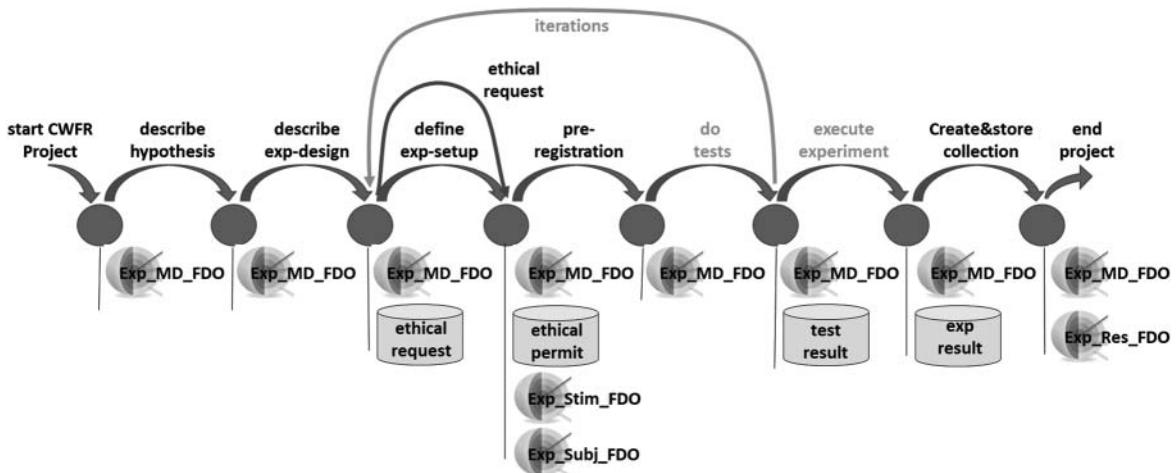


Figure 2. In a schematic way the actions and states within a typical workflow as introduced above are shown. The workflow consists of 9 canonical steps and indicates possible parallelism and iterations. The general glue that virtually integrates all information is the Exp_MD_FDO. It also contains references to all information being used in the process.

After the ethical percept has been received, some experimenters like to register the experiment to document it. If the registration site has an API, this step is also reduced to a simple automatic interaction where the needed information is extracted from Exp_MD_FDO. It should be noted that any repository that supports DOIP can be used to store FDOs. As always, it will be up to the implementations to generate indexes, etc. to support fast actions.

3.2 Test and Execution Phase

After all these preparations, tests can be carried out to ensure that the experimental paradigm and the chosen parameters together with the chosen hardware and software setup guarantee smooth experimentation as expected. Often these tests result in changing specific parameters and the list of stimuli, i.e., testing mostly results in iteration cycles (green arrow) to redefine the experimental setup. Sometimes even the experimental design and the hypothesis need to be adapted which would require starting the workflow steps again. This would result in a new Exp_MD_FDO object which can be instantiated based on the old one dependent on the step that is repeated, i.e., the researcher only needs to enter the changes.

To carry out the tests the experimental software that has been selected needs to be executed, i.e., a wrapper must be called receiving the Exp_MD_FDO object which includes all necessary information or references to the information needed. The wrapper then will transform the existing information in a way that the experimental software receives interpretable input. As indicated above, it would be helpful if the information in Exp_MD_FDO is structured and can be interpreted by machines which requires defined semantics registered in an open type registry which is a specific type of semantic artifacts, such as, provided by GWDG. At the end of the test runs the experimental software returns control to the wrapper which

creates an FDO called Exp_Res_FDO containing the usual metadata and a reference to the structured result file. This can be visualised with the means used by the researchers.

When the tests were not satisfying, going back to earlier steps will be required as indicated above. When they were positive, Exp_MD_FDO will be extended by a reference to Exp_Res_FDO to document successful testing and the execution software can be started again—now with all information (stimuli, subjects, parameters) for the real experiment. The result will be a new Exp_Res_FDO that contains all results in a structured form and Exp_MD_FDO will be updated to include the PID of the new Exp_Res_FDO.

3.3 Orchestration and Virtual Research Environment

Virtual Research Environments (VRE) (Candela et al., 2013) are designed to optimally support researchers in conducting the research while supporting orchestration and ensuring the produced data are managed according to RDM best practices (including FAIR Data Principles). During the orchestration, the sequence of steps needs to be specified. We assume the availability of components that can be chosen in a component software library, i.e., for each step indicated in the workflow as shown in Figure 3. Perhaps even more which have not yet been identified, a set of components should be available in the library.

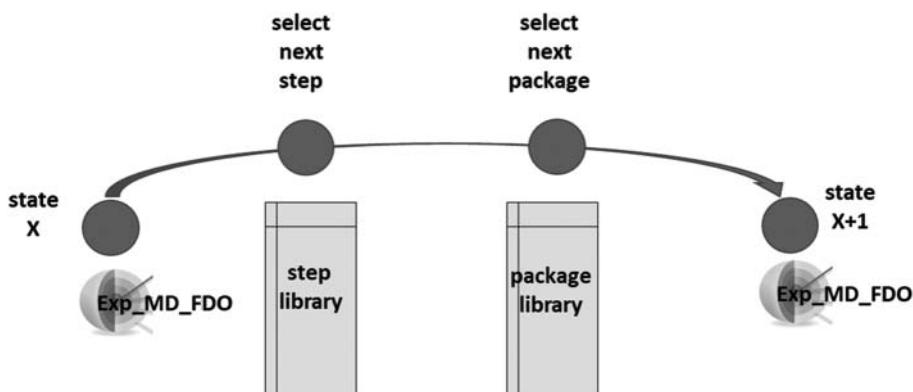


Figure 3. During orchestration, the researchers select useful components from a library that can help to move to the next state. The library will be organised offering canonical steps and specialised packages.

After each step a processing state X is being taken which is documented comprehensively by Exp_MD_FDO, i.e., all steps which have been chosen are described in a workflow process file (WPF) which adheres to some standard formats agreed upon by the broad workflow community and which is referred to from Exp_MD_FDO. WPF typically includes the sequence of steps and for each step a structure that can be used during execution to add process information. At state X , the VRE that supports orchestration will allow the user to select the software component to be launched to come to the next following state $X+1$. In general, this will happen in two actions to make it easy for the researcher: first a step is being chosen, and then a specific software package that addresses the needs. In the case of the experiment execution, even more filtering might be necessary to select the right software supporting the chosen paradigm.

We assume that different specialised libraries will emerge to support different research disciplines. We also assume that wrappers will be available where necessary to embed existing software packages. The glue that binds the workflow together is the Exp_MD_FDO. At each stage, it is updated to include all relevant information about the workflow process. We assume that either the software packages being integrated or the wrappers that embed the packages are supporting the Digital Object Interface Protocol (DOIP) and interact with a repository that is also able to talk DOIP independent of its local data organisation.

4. CONCRETE EXAMPLES

Specially created micro-workflows will be used to actually run the tests and the experiment using specific software packages. Many of them have been developed to meet specific requirements often with hard timing. A large variety of experimental paradigms are being used in the various experimental labs. Here we want to refer to examples that may indicate the differences.

4.1 Psycholinguistics

Often experiments are being carried out to get an idea about the pre-activation of a cohort of concepts when a certain target concept is being presented to a human subject (visually, acoustically). “Concepts” stored in the human mind are typically represented by activation patterns of a set of neurons. And these seem to be triggered by signals that are issued when other (semantically) related concepts are being recognised. The assumption is that when a specific item is being shown to a human subject, it will pre-activate a set of other items in the human mind with the result that if such a related item is being presented, subjects will respond faster.

The **Micro-Workflow** would then typically look like:

1. An image with an item is presented visually for a short time.
2. After some delay, a second image with another item is presented for a short time.
3. In parallel, a reaction timer is started at exactly the same moment.
4. The subject will speak loudly and mention the “name” of the item.
5. At speech onset the timer will be stopped, yielding the reaction time.
6. A record with all experimental parameters (image numbers, timing parameters, subject number, reaction time) is being added to a file.
7. This is iterated over several stimuli pairs and a set of subjects.
8. Finally, the final experimental result file is being stored safely.

The following aspects should be noted:

- There are different variants of such an experimental paradigm, i.e., reactions can be measured differently, time-series data (eye-tracking, brain imaging, etc.) could be measured in parallel to detect mental activity during stimulus processing, etc.

- Often special tools are being used to implement such Micro-Workflows in a way that narrow timing constraints are met. This implies that the CWFR shown in Figure 2 needs to be able to embed such micro-workflows at the steps “run tests”, “execute experiment”. It will be crucial to find a way to interface between the CWFR part and the Micro-Workflows since different tools are being used.
- Often these tools can't be changed so easily, i.e., they should not be loaded with the need to use or create digital objects. This should be done by the CWFR wrapper.

4.2 Experimental Economics

A large share of behavioural (economic) experiments are conducted in computerized laboratories, i.e., participants are placed in front of a computer where they receive information about the decision context. In addition, participants are informed about the available choices and the consequences of each choice. After that participants chose, either for themselves or in interaction with other participants, from these alternatives. Participants' behaviour or choices are usually recorded as inputs of numbers or text, or via different choice scales.

In most cases, decisions are stored anonymously. Meaning that personal information is not stored alongside the decision data because it is hardly used in data analysis. In case collecting demographic information is necessary, it is also stored in an anonymized or pseudonymized way [5]. The following steps describe a typical micro-workflow that can be found in most economic laboratory experiments.

Micro-workflow of a typical behavioural economic experiment:

- Participants receive an invitation to a particular experiment via e-mail. This e-mail contains information about the date, time, and length of the experiment.
- Participants who show up at the laboratory by the time of the experiment will be randomly seated at one of the computers.
- After a sufficient number of participants have arrived at the laboratory, the experimenter will present the instructions and rules for the experiment.
- In case participants interact with each other, they are usually randomly allocated to different groups. This group allocation is an important variable of the dataset.
- As soon as all participants signaled that they understood the instructions, the experiment is started.
- A large share of the experiments are played repeatedly, i.e., participants will face the same or a similar situation several times. All decisions are recorded in a spreadsheet.
- After the last decision has been made, participants may complete an additional questionnaire to describe their experience during the experiment, their motives, or their strategies.
- At the end of the experiment, participants are usually paid in cash or via online transfer. Their payment usually depends on their choices and/or the choices of other participants. The payment data is stored in a separate spreadsheet.

Note that the same workflow can be applied for online experiments or in field lab experiments. In the latter, data may also be recorded on paper and digitized later. The collection of behavioral data, as described above, may also be complemented by physiological measures, e.g., galvanic skin resistance (GSR), eye-tracking, or electroencephalography (EEG). The software that is used for behavioral economic experiments ranges from standardized questionnaire tools to very specific tools exclusively developed for economic experiments. As for the physiological measures, the hardware usually comes with a preconfigured set of software tools. These tools typically vary between different hardware providers.

In most cases the availability of tools and software packages is primarily affected by the laboratory, usually based on the demand of the most active researchers. The same is true for many requirements and regulations regarding data storage and data management. As a result of this situation, many laboratories store a large amount of raw data and auxiliary materials (like instructions, program code) on local devices. Typically, the data in these lakes is hardly described with any metadata besides the information that is necessary to organize the experiment (like date, time, number of participants). At the same time, these pieces of metadata are rarely transferred to published data sets.

All steps of the micro-workflow described above typically happen after a study design has been tested (Step 6 in Figure 1). Right after the actual experiment has been conducted (Step 7 in Figure 1) the anonymized data can be added to the existing FDO for this study.

4.3 Social Sciences

In the social sciences, such as sociology and political science, methodological paradigms in experimental research vary. For example, the methodology of the behavioral game theory having its roots in Experimental Economics as described above, which follows the assumptions of induced value theory [6], has been adopted by other disciplines like the social sciences. Researchers following this experimental paradigm use similar workflows and metadata across disciplines. Machine-readability is thereby easy to achieve in this context, as the range of core metadata, for example, on experimental design as well as on experimental setup, is predetermined to a large extent. As a result, the tools used in game theory experiments hold a critical position for CWFR and FDOs as a whole. To fully realize the potential of CWFR, the MD_FDO describing the experimental design must reflect the degree to which it meets the standards of game theory experiments. For example, the no-deception clause and an appropriately high monetary payoff are two mandatory properties to be observed. Deviation from these standards, e.g., identical payoff levels for students and professionals, could cause misinterpretation of the results.

The steps described in Figure 1 are consistent with the methodological paradigm used in an experiment. However, good scientific practice is not guided by a particular method but by discipline. In the field of analytical-empirical sociology, step 5 is becoming increasingly recommended (pre-registration) [7]. In experimental economics, according to the designated standards of the Society for Experimental Economic Research (GfeW), step 5 can be omitted if full reporting is provided as an alternative.

Survey experiments, e.g., factorial surveys [8] or conjoint analyses [9], measure attitudes and behavioral intentions rather than behavior. If reusability of data across methodological research paradigms is intended, these assumptions need to be documented in the MD_FDO of the data, which will be published in the final step of the project. Sociology also involves experiments with behavioral interventions, which have their methodological background in social psychology. In addition, clinical randomized controlled trials in the context of medical sociology might include studies with “experimental” treatments compared to “control” treatments and control groups without intervention. However, regardless of methodological paradigm or discipline, all types of experiments follow the canonical steps described above.

4.4 Medical Domain

In the case of medical experiments (especially randomized controlled trials, RCTs), in-depth quality assurance procedures must be installed in most work steps and the applicable quality guidelines and standards must be adhered to. These include the GxP standards known from many related fields, e.g., the GCP—Good Clinical Practice Guide or the GAMP 5 Guide. The requirements for electronic systems for processing study data are manifold and also cover the workflows discussed here. However, since the entire system must always be validated and certified, the workflows as part of the system (and also the CWFR in general, at least if they have been implemented in advance) can generally be assessed as meeting the requirements, and a high level of quality and fluidity of work is ensured. Also, the GAMP 5 guide, like all other GxP requirements, requires compliant computerized system validation. This is a requirement that should not have to wait much longer for legislative implementation.

Speaking of the law, as soon as the data to be processed is patient data, and thus Patient Health Records (PHR), the strictly regulated terrain of medical devices is quickly entered into. Exemplary, but truly not exhaustive, are the harmonized ISO 13485:2016 and ISO 62304:2016, which define the requirements for software as a medical device (SaMD) with regard to regulatory purposes and control product compliance. Requirements for interoperability have also now found their way into digital medicine: in addition to the Medical Devices Regulation (MDR 2017/745), the (German) standards of Sections 394 ff. of the German Social Code, Book V (SGB V) also stipulate that information technology systems must be semantically, syntactically and structurally compatible. The demanding area of data protection and data security needs no mention; here, the immense requirements should be generally known.

In addition to this brief excursion into the regulated world of medical devices, there is a large number of recognized standards that frame medical research, research software systems, and also workflows. FAIR principles have also begun their triumphant march in medical research. Driven by the FAIR consortium of representatives from academia, industry, funding agencies, and scientific publishers, which can be described as a growing “data science community,” existing and new data are being discovered, integrated, and analyzed. For example, at the metadata level, to develop a FAIR registry for medical data, taking into account the FAIR Data Principles. It is here that the demand for interoperability, functioning workflows, and secondary use in research becomes clearly audible.



5. CONCLUSIONS

Analysing experimental practices in laboratories across different disciplines collecting behavioural data from human participants indicates that in all these places similar steps are being taken which we can indeed call canonical steps in workflow scenarios. Some of these steps have different flavours, such as for the ethical review where all research organisations have their own templates. However, requests with some exceptions, as in medical laboratories, the same kind of information from the researchers is needed across the disciplines and methodological paradigms. Therefore, it would make sense to develop a CWFR like framework across disciplines, provide libraries of packages organised by steps, and allow researchers to easily adapt canonical workflows to their specific needs which then can be executed repeatedly. These would guide them through the various steps without the need to be bothered by details.

These workflows will need to embed existing tools, such as, for selecting the set of subjects and stimuli that will be used for executing the experiments. The different labs have developed their own tools for these steps partly integrated with databases which would have to be embedded by wrappers also doing the conversion of the gathered parameters so that these external programs can use them as input. Especially, the experiment execution software packages are often highly specialised due to the hardware being controlled (special instruments), the tight timing tolerances, and the embedded micro-sequences of actions. Obviously, the software for these tool integrations by wrappers would have to be developed by experts.

Experiments with humans imply a number of special features which a workflow machinery needs to address:

- Experimental designs and parameters are usually revised at the beginning of the research process which may require many iterations until the final settings have been established.
- Experimental workflow execution is highly asynchronous since at some steps researchers need to wait on external signals or apply parallelisms.
- Often experiments are being executed in different labs due to access to different instruments and different types of subjects.

These requirements and the recurring patterns in these experiments suggest an implementation of the CWFR principles and base all experiment documentation on FDOs being hosted in any DOIP adapted repository for all steps assuming that the structure of the content has been standardised. It would take off administrative work from the researcher, widely improve the efficiency of the experimental work, create FAIR compliant documentation of all experimental steps without bothering the researchers, and would make it easily possible to run experiments in different labs.

AUTHOR CONTRIBUTIONS

All authors contributed ideas, text, and review comments in the production of the paper. C. Blanchi (cblanchi@cnri.reston.va.us) is a core developer of components of the Digital Object Architecture and therefore a major contributor to the paper. P. Wittenburg (peter.wittenburg@mpcdf.mpg.de) is co-editor of the FAIR principles paper and member of the FDO Forum and a major contributor to the FAIR and FDO aspects in the production of the paper.

REFERENCES

- [1] Wilkinson, M.D., et al.: The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3, 160018 (2016)
- [2] Jeffery, K., et al.: Not ready for convergence in data infrastructures. *Data Intelligence* 3(1), 116–135 (2021)
- [3] Wittenburg, P., Strawn, G.: Common patterns in revolutionary infrastructures and data. B2Share. Available at: <https://b2share.eudat.eu/records/4e8ac36c0dd343da81fd9e83e72805a0>. Accessed 17 December 2021
- [4] De Smedt, K., Koureas, D., Wittenburg, P.: FAIR digital objects for science: From data pieces to actionable knowledge units. *Publications* 8(2), (2020)
- [5] Friedman, D., Sunder, S.: *Experimental methods: A primer for economists*. Cambridge University Press, Cambridge (1994)
- [6] Smith, V.L.: Experimental economics: Induced value theory. *The American Economic Review* 66(2), 274–279 (1976)
- [7] MZES Open Social Science Conference (2019). MZES-GESIS Pre-Registration Challenge. Available at: <https://www.mzes.uni-mannheim.de/openscience/index.php/pre-registration-challenge/>. Accessed 17 December 2021
- [8] Auspurg, K., Hinz, T.: Factorial survey experiments (Vol. 175). Sage Publications, Los Angeles (2014)
- [9] Green, P.E., Krieger, A.M., Wind, Y.: Thirty years of conjoint analysis: Reflections and prospects. *Interfaces*, 31(3_supplement), S56–S73 (2001)

AUTHOR BIOGRAPHY

Dirk Betz is a Senior Researcher at TIB-Leibniz Information Centre for Science and Technology Hannover, affiliated with the Center for Empirical Research in Economics and Behavioral Sciences (CEREB) at the University of Erfurt, and a co-spokesperson of the METHODS NFDI-consortium project. He received his Ph.D degree from Cologne University, Germany. Together with Claudia Biniossek, he developed the repositories x-econ and x-science for data from experimental economics and social sciences.

ORCID: 0000-0002-6411-4758



Claudia Biniossek is a Senior Researcher at TIB-Leibniz Information Centre for Science and Technology Hannover, and co-spokesperson of the METHODS NFDI-consortium project. Claudia Biniossek is affiliated with the Center for Empirical Research in Economics and Behavioral Sciences (CEREB) at the University of Erfurt. Her research interests lie in bringing together theories, methods, and data infrastructures in a transdisciplinary approach to improve research on human behavior. Together with Dirk Betz, she developed and established the repositories for behavioral economics (x-econ.org) and experimental social sciences (x-science.org).

ORCID: 0000-0002-2202-7875



Christophe Blanchi is the Executive Director of the DONA Foundation in Geneva where he is responsible for its day-to-day operations, promoting and evolving the Digital Object Architecture and its related set of standards, and insuring the consistent operations of the Global Handle Registry. Prior to joining the DONA Foundation, Christophe Blanchi was a senior research scientist at CNRI in Reston, Virginia, where he was involved in research, development, and deployment of Digital Object Architecture related technologies.

ORCID: 0000-0003-2277-5176



Felix Henninger is lead developer of the lab.js data collection framework for distributed online behavioural, cognitive and neuroscientific research. He is a Ph.D. candidate at the Graduate School of Economic and Social Sciences (GESS) at the University of Mannheim, from which he holds a Master's degree in Psychology. He is an avid advocate for Open Science, having initiated the Open Scholarship Knowledge Base in collaboration with the Center for Open Science, a founding member of the German Reproducibility Network, co-organiser of the Mannheim Open Science Meetup, and co-founder of the Landau Open Science Working Group.

ORCID: 0000-0002-7730-9511



Thomas Lauer is a behavioral economist with a Ph.D from the University of Erfurt (2011). He works as the scientific managing director at the Erfurt Laboratory for Empirical Research. His research is mainly focused on how people's expectations, motives and preferences affect their behavior in strategic contexts, and how the institutional mechanisms of these contexts shape people's choices. One of the main questions addressed in his research is how cooperation among people can be fostered and what organizational conditions are necessary to sustain this cooperation. As a spokesperson of the Erfurt Open Science Initiative he also works to institutionalize open science practices and helps to establish structures to aid researchers in opening up their research processes.

ORCID: 0000-0002-7662-1132



Philipp Wieder received his Ph.D degree from TU Dortmund, Germany. He is the Deputy Head of the Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), Germany, and is teaching Data Science Infrastructures at the Georg-August-Universität Göttingen, Germany. He is contributing to areas like distributed infrastructures, research data management, and data analytics for several years. His research interest lies in distributed systems, service level agreements, and resource scheduling. He is actively involved in the German National Research Data Initiative (NFDI) and several projects related to the European Open Science Cloud.

ORCID: 0000-0002-6992-1866



Peter Wittenburg has a background in electrical engineering, has been working as Technical Director at the Max Planck Institute for Psycholinguistics for many years and acted as member of the IT Advisory board of the president of the Max Planck Society (MPS). The Max Planck Institute was from the beginning focusing on digital technologies to understand the functioning of the brain with respect to language processing. The institutes need for getting access to data from other institutes to feed the stochastic engines they applied rather early led him to become an expert in building data/research infrastructures. He was responsible for the technological aspects of three large international and European research infrastructures: DOBES, CLARIN and EUDAT. In this function he understood that data work across silos is highly inefficient and that harmonisation and standardisation is required to improve the situation. This was the reason that he co-founded the Research Data Alliance in 2013 and the FAIR Digital Object Forum in 2019.

ORCID: 0000-0003-3538-0106



Martin Zünkeler is founder and managing director of KAIROS, a spin-off of Charité Berlin, which develops and offers solutions in the field of biobanking and medical informatics for scientific research institutions and university hospitals. He is particularly interested in bringing together different healthcare stakeholders to jointly develop efficient solutions. Martin Zünkeler strongly believes that data exchange can only succeed through workflows. Since 2003, he has been involved in bringing open-source software workflow engines to the healthcare market to orchestrate different steps of data capture. Martin Zünkeler aims to achieve his goal of linking *in vivo*, *in vitro*, and *in silico* technologies to take precision medicine to the next necessary level.

ORCID: 0000-0002-2098-8779